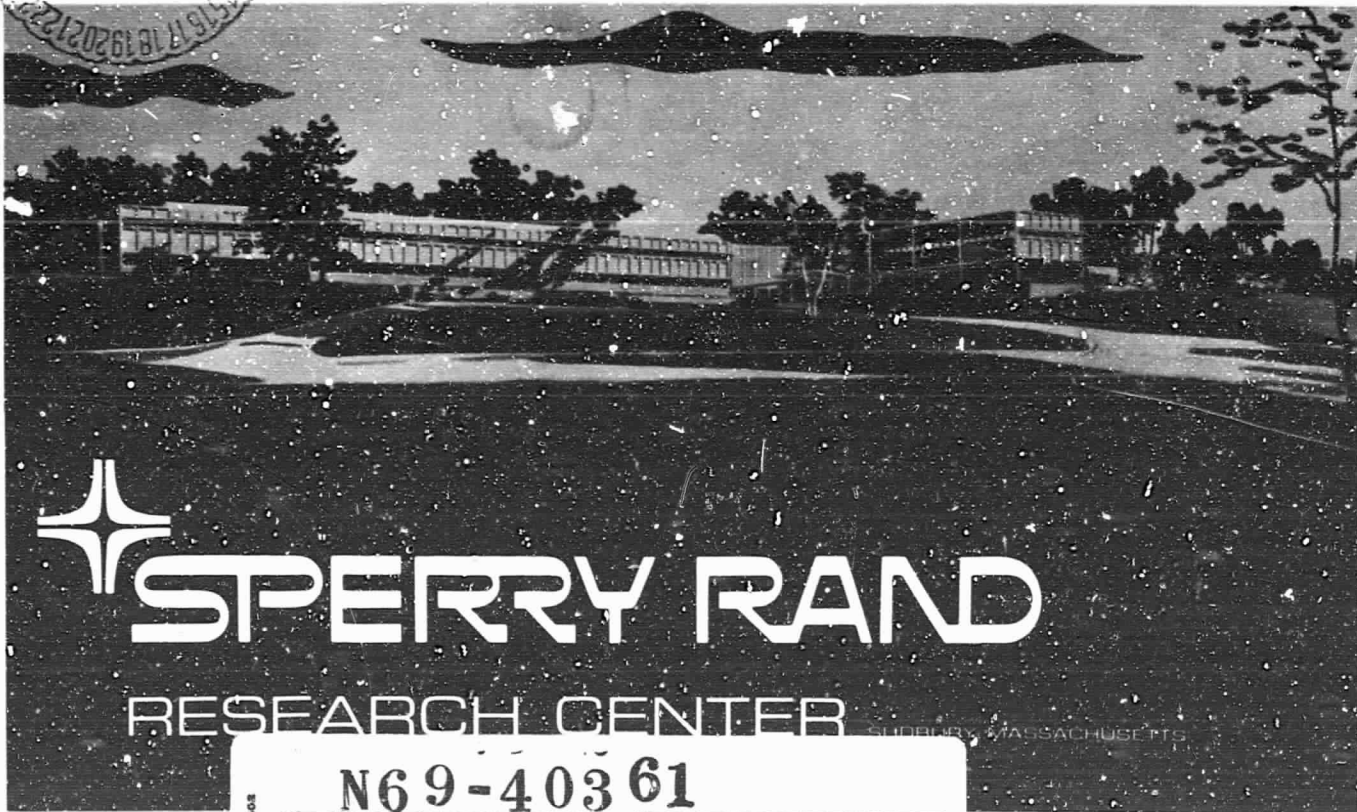
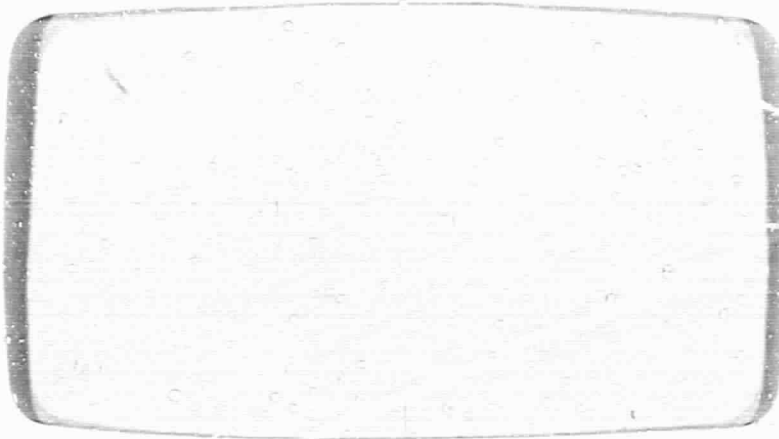


General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.
- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.
- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.
- This document is paginated as submitted by the original source.
- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

CR 86251



SPERRY RAND

RESEARCH CENTER

SUDBURY, MASSACHUSETTS

FACILITY FORM 502

N69-40361

(ACCESSION NUMBER)

22

(PAGES)

CR-86251

(NASA CR OR TMX OR AD NUMBER)

(THRU)

1

(CODE)

09

(CATEGORY)

SRRC-CR-68-20
April 1968

THIRD QUARTERLY TECHNICAL REPORT

INVESTIGATION OF NEW CONCEPTS
OF ADAPTIVE DEVICES

Contract No. NAS 12-570

Period Covered:
15 December 1967 - 14 March 1968

E. Lewis, M. Cohn, A. Lincoln, F. Sewell
M. Doig, H.A.R. Wegener
Sperry Rand Research Center, Sudbury, Massachusetts

Prepared for
NASA Electronics Research Center
Cambridge, Massachusetts

I. INTRODUCTION

This quarterly report will discuss an application of the MNS variable threshold transistor, a transient solution of the charging current, some elements of device fabrication, and a summary of a technique for the predication of charge storage at the I_2 - I_1 interface.

For the application of the memory transistor a "learning machine" was constructed that utilizes the charge stored in the gate structure to control the threshold voltage that, in turn, can be used to influence a series of logic functions. The output of these logic functions consists of a "reward" and a "punish" decision. It is shown that learning can be accomplished even by an imperfect "trainer". The only requirement of the trainer is that he make more correct decisions than incorrect decisions.

The transient solution for the charging current in a MI_2I_1S structure has been accomplished by an analog computer. The analog computer has solved a differential equation that was formulated from an electrical equivalent circuit of the device structure. This result compares favorably with an approximate mathematical solution.

The remainder of this report presents some processing results and a summary of the method of charge storage prediction. It is demonstrated that a thermal oxide interface is more desirable than one grown pyrolytically.

II. A MATHEMATICAL MODEL OF "STUDENT" TYPE DEVICES

A. Introductory Remarks

An electronic device, "Student-001", has recently been constructed

at SRRC in order to demonstrate the use of MNS variable threshold transistors in circuits which appear to learn from experience. Upon being stimulated by the flipping of a switch, the device responds by lighting or not lighting a lamp. The device is then rewarded or punished (through switches) in order to encourage or discourage its most recent response.

Although no systematic experiments have been performed, several users of the devices have been convinced that Student-001 learns, statistically speaking, from its training. The object of this report is to provide a mathematical model for the device and its immediate logical extensions whereby their behavior may be predicted quantitatively, and experiments may be designed and further extensions planned.

B. The Single-Light Student

A single-light student's light is controlled by a variable threshold transistor whose input voltage, v_i , is randomly generated by a sampling circuit (activated by the "stimulate" switch) from a uniform distribution between v_{il} and v_{iu} . The light is turned on if the input voltage is exceeded by the threshold voltage v_t , that is, if $v_i < v_t$. The threshold voltage is altered by the "reward" and "punish" switches in such a direction as to raise or lower the probability of repeating the most recent response.

For instance, if the desired condition is "on", then an "on" response will lead to "reward", which raises v_t , say by an amount $\Delta^+ v_t$; if the response was "off", then the "punish" switch would likewise raise v_t by $\Delta^+ v_t$. Conversely, if the desired condition is "off", then responses of "on" and "off" will lead respectively to punishment and reward, both of which will lower v_t by an amount $\Delta^- v_t$. The threshold voltage ranges between upper and lower limits, v_{tu} and v_{tl} .

These may be adjusted to lie outside, on, or inside the corresponding input voltage limits; suppose for the present that they coincide, that is, $v_{tu} = v_{iu}$, and $v_{tl} = v_{il}$; suppose also that $\Delta^+ v_t = \Delta^- v_t = \Delta v_t$.

Let the instantaneous state of the student be described by the probability that at the next stimulus the light will go on, let the goal variable g be defined by

$$\begin{aligned} g &= 0 \quad \text{if desired response is "off"} \\ &= \quad " \quad " \quad " \quad " \quad \text{"on"}, \end{aligned}$$

and let $m = \frac{v_{iu} - v_{il}}{\Delta v_t}$. Here v_t is being quantized to m discrete values. But since

$$\begin{aligned} p &= \Pr\{v_i < v_t\} \\ &= \Pr\{v_{il} \leq v_i < v_t\}, \end{aligned} \tag{1}$$

and by the assumption of the uniform distribution of v_i over the interval $[v_{il}, v_{iu}]$,

$$p = \frac{v_t - v_{il}}{v_{iu} - v_{il}}. \tag{2}$$

Thus p also is being quantized into m intervals, each defining a state of the Student.*

* A continuous analysis of student could certainly be carried out. The quantization, however, is convenient for the sake of exposition, especially in the multi-dimensional case.

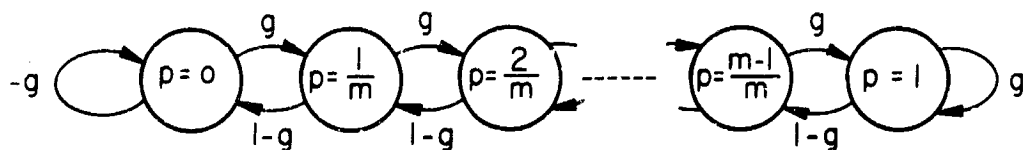


FIG. 1(a) State diagram of a one-dimensional student.

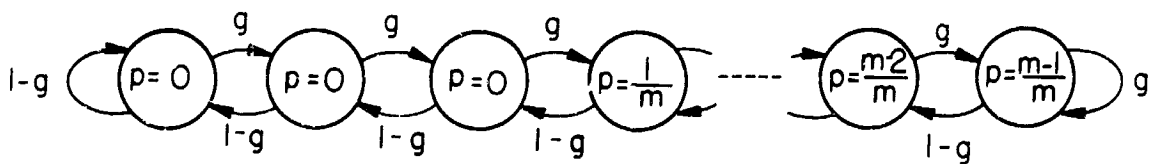


FIG. 1(b) Student of Fig. 1(a) with unequal input and threshold voltage limits.

A state diagram of a one-dimensional Student is shown in Fig. 1a.

Figure 1b depicts the same Student but with

$$v_{tu} = v_{iu} - \Delta v_t \quad (3)$$

and

$$v_{tl} = v_{il} - 2\Delta v_t . \quad (4)$$

It is clear that the state-diagram structure remains essentially the same. Only the number of states and their labels are altered when the equality of voltage limits is relaxed.

Another useful description of these processes is the transition matrix shown in Fig. 2.

$$M = \begin{vmatrix} 1-g & g & 0 & 0 & 0 \\ 1-g & 0 & g & 0 & 0 \\ 0 & 1-g & 0 & g & 0 \\ 0 & 0 & 1-g & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-g \\ 0 & 0 & 0 & 0 & 0 \end{vmatrix}$$

Fig. 2. Transition Matrix for State Diagram Structures of Figs. 1a, b.

Since all entries are non-negative, and the row sums are unity, this matrix describes a Stochastic (or Markov) process. For g , either 0 or 1, the process has an absorbing state (either the first or last), and is connected to this state in the sense that it can be reached from any other. These conditions are sufficient to confirm mathematically the physical intuition that if g is held

fixed for a sufficiently long time (at most m trials), the process will enter and remain in the intended stable or absorbing state. In other words, the Student will learn in m or fewer lessons. Obviously the fastest learning occurs when $m=1$.

Stochastic Training

The foregoing analysis has dealt with a perfect, or deterministic trainer. It is also instructive to consider an imperfect, or stochastic trainer, who sometimes errs in his reward or punishment of the Student. The simplest instance of such a trainer is one who errs with probability $1-q$, independently of his history (and of the Student's state). In terms of the transition matrix M of Fig. 2, in training the light to go on, the trainer at each trial chooses $g = 1$ with probability q .

It can be shown algebraically that for the purpose of this analysis that this is equivalent to substituting $g = q$ in the transition matrix. Heuristically, this is explained by saying that instead of flipping a biased coin to determine which deterministic transition matrix is to apply to the Student, the trainer lets the Student flip the coin and make the appropriate transition.

With $g = q$, it can be confirmed that the stationary distribution vector \underline{s} of M , that is, the unique state occupancy distribution \underline{s} , is such that

$$\underline{s}M = \underline{s} \quad ,$$

is given by

$$\underline{s} = \frac{1-r}{1-r^{m+1}} \left(r^m, r^{m-1}, \dots, r, 1 \right) , \quad r = \left(\frac{1-q}{q} \right). \quad (5)$$

This means that after many trials the Student achieves a geometric state occupancy distribution, whose maximum occurs in the last state (the desired one, where the probability of lighting is $p_m = \frac{m}{m} = 1$), with a value there of

$$\frac{1-r}{1-r^{m+1}}, \quad r = \left(\frac{1-q}{q} \right). \quad (6)$$

The probability, after many trials, of the Student responding correctly is found by summing the correct response probabilities in all the states, weighted by the state occupancy probabilities:

$$\begin{aligned} \bar{p} &= \sum_{i=0}^m p_i s_i = \sum_{i=0}^m \frac{i}{m} \frac{(1-r)r^{m-i}}{(1-r^{m+1})} \\ &= \frac{m-(1+m)r+r^{m+1}}{m(1-r)(1-r^{m+1})}. \end{aligned} \quad (7)$$

If $q > \frac{1}{2}$, then $r < 1$, and the probability \bar{p} of correct response approaches 1 as m grows large. This implies that by choosing the threshold voltage increment size sufficiently small, the Student will learn to give the correct response with probability arbitrarily close to 1 after many trials, despite the imperfect trainer. (The trainer, however, may not be so imperfect as to train incorrectly more often than not.) The formula given above for \bar{p} can be used to find the Student's ultimate probability of correct response for arbitrary voltage quantization and trainer error rate.

C. Electrical Description of "Student"

The one-light Student, Fig. 3, comprises three sections. These are

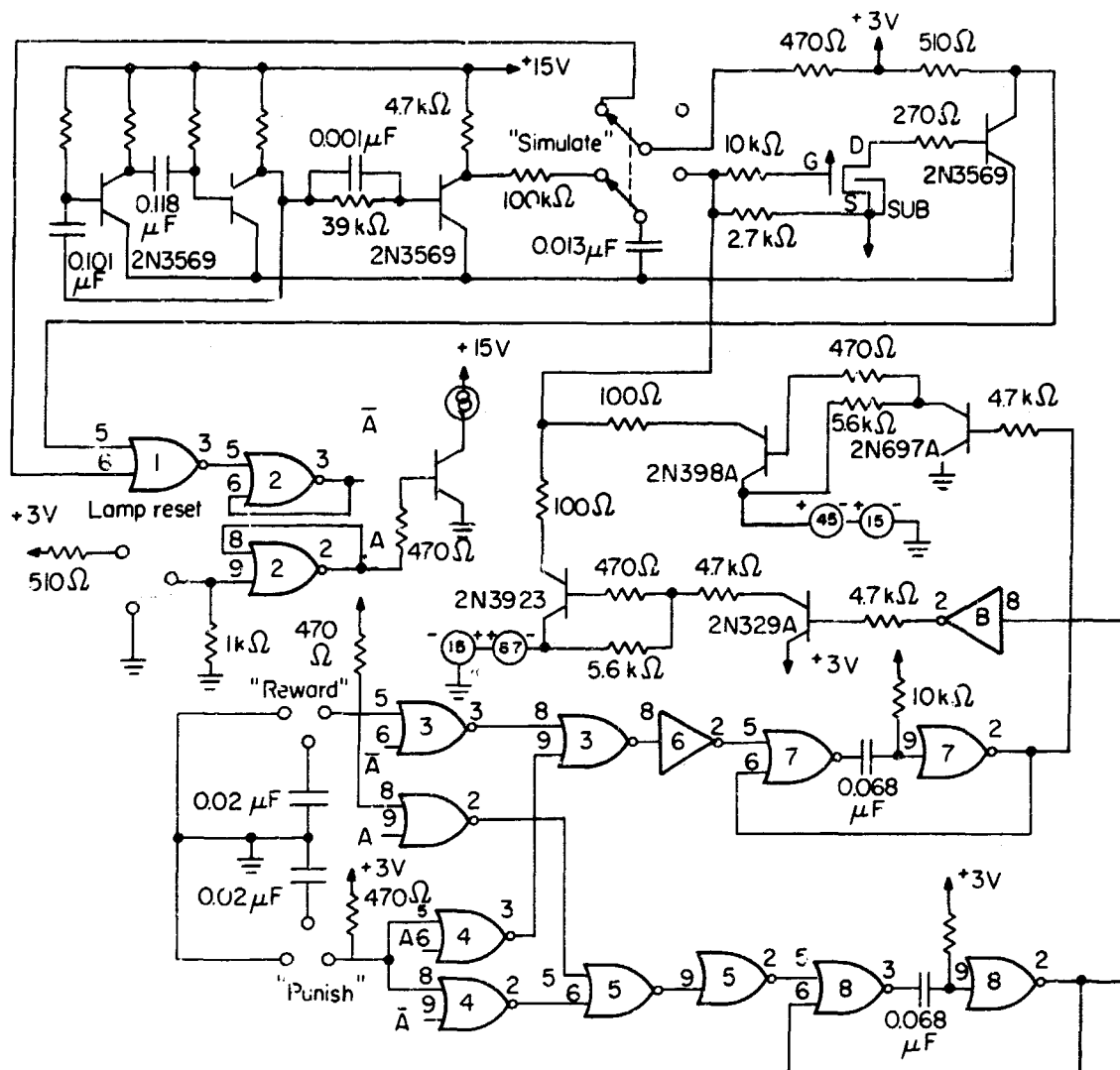


FIG. 3 Student 1-001 MNS-VTT learning system.

the stimulus generator, the threshold storage and response unit, and the conditioning unit. The stimulus generator consists of a free-running multi-vibrator with a frequency of approximately 1 kHz whose output is clipped and used to drive a series RC integrating circuits. The capacitor in the integrating circuit is connected to the resistor through a double-throw switch marked "STIMULATE". When this switch is thrown the voltage appearing across the capacitor is held and transferred to the input of the threshold storage unit.

The threshold storage and response unit contains an MNS-VTT at its input. This element stores the present learning threshold in its gate insulator region. If the voltage transferred to its gate by the stimulus switch is more negative than the stored threshold, the MNS-VTT conducts and sets a flip-flop. The output of the flip-flop drives a lamp. Conduction of the MNS-VTT is indicated by illumination of the lamp. A second transfer contact on the "STIMULATE" switch is used to prevent the flip-flop from being set when the voltage-transfer capacitor is not connected to the MNS-VTT gate. The flip-flop is reset by momentarily activating its resetting input by means of the "reset" switch.

The conditioning unit contains logic elements which determine whether to apply a positive or a negative pulse to the gate of the MNS-VTT, based upon whether an impulse is received from the "reward" or "punish" switches and upon the state of the flip-flop. These pulses are shaped by single-shot multivibrators and used to drive the common-output, high-voltage, pulse generators which connect to the MNS-VTT gate circuit.

Power is supplied to Student by internal batteries and is controlled by the two power switches on the front panel.

III. ANALOG COMPUTER SOLUTION FOR THE CHARGING CURRENT IN AN MI_2I_1S CAPACITOR

A. Device Transient Measurements and Model Correlation

The MI_2I_1S capacitor can be represented by the simplified electrical equivalent shown in Fig. 4 . If the device is pulsed with a continuous chain of pulses with the same polarity, then a steady-state situation is reached where the interface between I_1 and I_2 can no longer accept charge. When this condition has been established, the current generator, i_1 , becomes inoperative, i.e., $i_1 = 0$. For this special case the electrical equivalent may be represented as shown in Fig. 5.

For the case where i_1 is operative, the solution for i_1 as a function of the measurable voltage v_D is

$$i_1 = C_1 \left(1 + \frac{C_o}{C_I} \right) \frac{dv_D}{dt} + \frac{v_D}{RC_I} - \frac{dV_{appl}}{dt} , \quad (8)$$

where,

$$C_I = \frac{C_1 C_2}{C_1 + C_2} .$$

When i_1 is no longer operative, Eq.(8) reduces to

$$\frac{dV_{appl}}{dt} = \left(1 + \frac{C_o}{C_I} \right) \frac{dv_D}{dt} + \frac{v_D}{RC_I} . \quad (9)$$

In order to minimize the uncertainties in the response of this device, the applied voltage pulse train consisted of well defined ramp functions and the output capacitance was deliberately increased by 100 picofarads. The solution for v_D with a

ramp excitation is,

$$v_D = KRC_I \left(1 - e^{-\frac{t}{R(C_o + C_I)}} \right) \quad (10)$$

where,

K is the slope of the ramp excitation

$$K = -0.33 \times 10^6 \text{ V/sec}$$

$$R = 10^5 \text{ ohms}$$

The measured response is shown in Fig. 6. The response shown in this figure can be represented by the equation

$$v_D = -0.28 \left(1 - e^{-\frac{t}{13.6 \times 10^{-8}}} \right) \quad (11)$$

Since R and K are relatively well known it is then possible to determine C_I and C_o using Eqs.(10) and (11).

$$C_I = \frac{-0.28}{KR} = 8.5 \times 10^{-12} \text{ farads} \quad (12)$$

$$C_o = \frac{13.6 \times 10^{-8}}{R} - C_I = 127.5 \times 10^{-12} \text{ farads} .$$

The value calculated for C_I corresponds identically to that measured using a capacitance bridge. The value for C_o is higher than the external 100 pico-farads since it also includes probe capacitance and additional interface capacitance from external circuitry. The primary reason for this calculation is that it establishes the fixed parameters in the model.

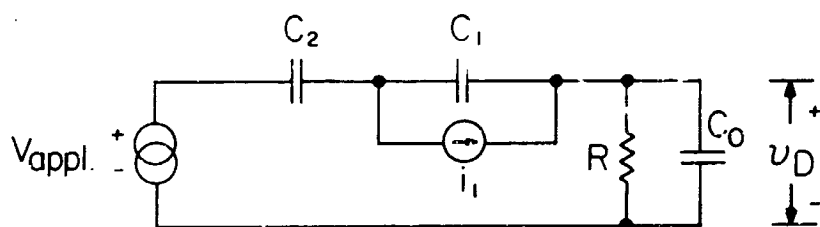


FIG. 4 Electrical equivalent of MI_2I_1S capacitor.

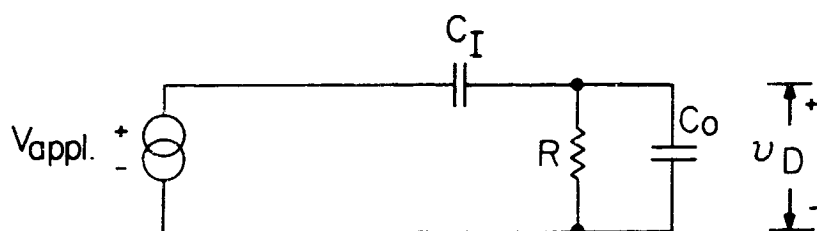


FIG. 5 Electrical equivalent of MI_2I_1S capacitor with i_1 dormant.

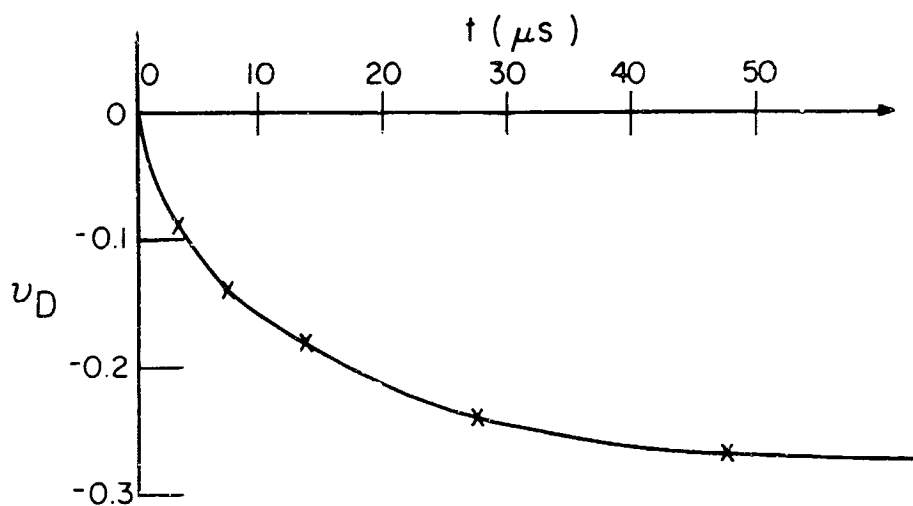


FIG. 6 Measured $v_D(t)$ with a negative ramp input.

B. Transient Measurement with i_1 Operative

By resetting the device with a positive pulse before the application of the negative ramp, it is then possible to observe the effect i_1 has on the measured response v_D . This is shown in Fig. 7. By approximating this response by two linear equations it is possible to calculate i_1 and then to compare it with an analog computer solution. The analog computer was described in the last quarterly report. For $t < 24 \mu s$, v_D may be approximated by

$$v_D \approx -1.13 \times 10^{-4} t. \quad (13)$$

and for $t > 24 \mu s$,

$$v_D \approx -0.27 \text{ V}. \quad (14)$$

Using these approximations for v_D , i_1 may be calculated using Eq. (8).

$$i_1 \approx 3.06 \times 10^{-6} - 0.266t \quad (\text{for } t < 24 \mu s) \quad (15)$$

$$i_1 \approx -0.26 \times 10^{-6} \text{ for } t > 24 \mu s. \quad (16)$$

A plot of these results is shown in Fig. 8. The $i_1(t)$, as calculated by the analog computer is also shown in this figure. There is a discrepancy between the two curves for $t > 24 \mu s$. This could be due to the effect of the approximation made for the mathematical description of v_D and the inexactness of the equivalent components used in the analog computer.

C. Conclusions and Further Study

It appears that the two layer representation of the MI_2I_1S capacitor

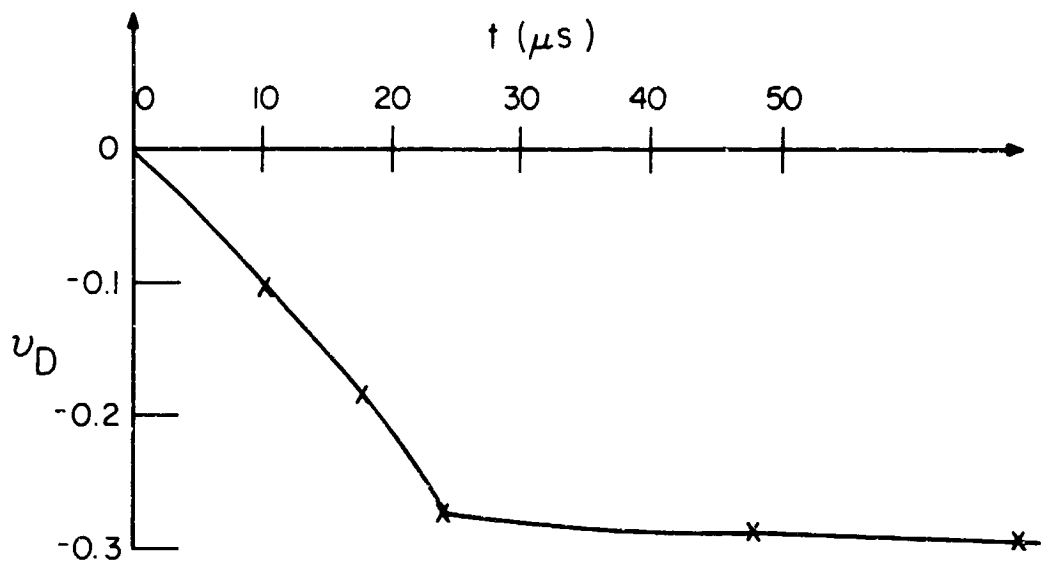


FIG. 7 Response of device to a negative ramp with a positive reset pulse.

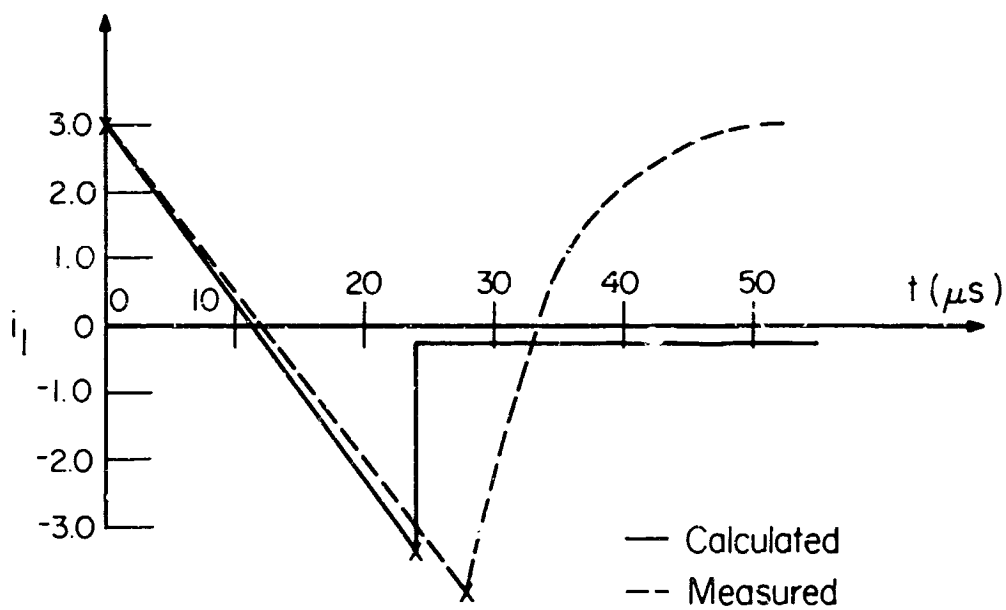


FIG. 8 Calculated and measured $i_1(t)$.

as shown in Fig. 4 is valid. By integrating $i_1(t)$ with respect to time, it is possible to obtain the charge transferred into the insulator interface and relate this to an offset voltage. The offset voltage calculated in this manner has been compared to that determined by a C-V plot and the two values are in good agreement.

The analog computer needs further refinement and adaptability features to facilitate component changes to correspond to device parameter changes. It is also possible that semiconductor space charge effects are influencing the device response. If this is the case then special measurement techniques must be devised that will eliminate this nonlinear dependence. The dependence of the device response on duty cycle variations and pulse magnitude variations needs further study.

IV. DEVICE FABRICATION

Two sets of memory transistor slices were processed using silicon nitride deposited in a hot wall furnace at 750°C . Since work in a cold wall system indicated that a very thin oxide at the interface gave lower threshold voltages and more reproducible results, slices were run in the hot wall with pyrolytic and thermal oxide interfaces. The gate structure was the same for both sets; 50 Å of oxide, 200 Å of silicon nitride and 800 Å of oxynitride.

The set of transistors processed with the pyrolytic oxide interface gave the following results when subjected to a 10 msec positive and negative 50 V pulse.

<u>Lot No.</u>	<u>Original V_T</u>	<u>V_T (+50 volt)</u>	<u>V_T (-50 volt)</u>
L-6-1	-8.0	-5.5	-6.0
L-12-1	-4.0	+2.0	-2.0
L-12-2	-4.0	-2.2	-4.0
L-14-1	-10.0	-10.0	-9.0
L-16-1	-10.0	-5.0	-9.0

These results indicate that there is a sufficient oxide thickness at the interface to lower the threshold voltages, but a thinner oxide interface is necessary to maintain a maximum total shift on pulsing.

The second set of transistors was processed with a thermal oxide interface to give the following results.

<u>Lot No.</u>	<u>Original V_T</u>	<u>V_T (+50 volt)</u>	<u>V_T (-50 volt)</u>
L-16-2	-6.0	-4.5	-15.0
L-18-1	-2.0	+2.0	-11.0
L-18-2	-1.0	+8.0	-11.0
L-21-1	-4.5	-8.0	-11.0

This set shows more reproducibility and better shifting characteristics.

Further work will be continued with thermal oxide interface in the hot wall furnace as well as the cold wall system. Additional work will also be done to obtain more processing control of the interface prior to nitride depositions.

V. CHARGE STORAGE PREDICTION FROM J-E CHARACTERISTICS

In this section we will discuss an approach showing that the storage-bias behavior (V_{FB} vs V_{appl}) for an MI_1I_2S device can be approximated without a detailed knowledge of the individual J-E curves, so long as one layer can be assumed to be much more conductive than the other.

The physical principles which govern the device performance are Kirchhoff's law, displacement vector continuity, and current continuity:

$$E_2(d-x) + E_1x = V_{appl} \quad (17)$$

$$E_2\epsilon_2 - E_1\epsilon_1 - \sigma = 0 \quad (18)$$

$$E_2\gamma_2(E_2) - E_1\gamma_1(E_1) = 0 \quad (19)$$

where x and $d-x$ are the thicknesses of layers I_1 (contiguous with the silicon), and I_2 (contiguous with the metal electrode) and ϵ_1 and ϵ_2 are the respective permittivities. The symbols $\gamma_1(E_1)$ and $\gamma_2(E_2)$ are the functional conductivities of each layer. Even though $\gamma_1(E_1)$ and $\gamma_2(E_2)$ are not known the equations may still be solved for E_1 , E_2 and σ (the charge density stored at the interface):

$$E_1 = \frac{V\gamma_2(E_2)}{x\gamma_2(E_2) + (d-x)\gamma_1(E_1)} \quad (20)$$

$$E_2 = \frac{V\gamma_1(E_1)}{x\gamma_2(E_2) + (d-x)\gamma_1(E_1)} \quad (21)$$

$$\sigma = \frac{V[\gamma_1(E_1)\epsilon_2 - \gamma_2(E_2)\epsilon_1]}{x\gamma_2(E_2) + (d-x)\gamma_1(E_1)} \quad (22)$$

These are the fields and the charge that will obtain when a voltage V_{appl} has been applied to a MI_1I_2S system for a time long enough so that equilibrium is established. The equilibrium value of these quantities is independent of any charge that may initially have been stored in the device.

Let us now assume that $\gamma_1(E_1) \gg \gamma_2(E_2)$ for any E_1 and E_2 and that $d-x \approx x$. Then Eq. (22) becomes

$$\sigma = \frac{V_{\text{appl}}\epsilon_2}{d-x} \quad (23)$$

But the flat band voltage (V_{FB}) which can easily be measured, is related to σ by $V_{\text{FB}} = \sigma \frac{d-x}{\epsilon_2}$. Therefore, Eq. (23) becomes

$$V_{\text{FB}} = V_{\text{appl}} \quad (24)$$

Thus, in this approximation, the storage-bias relation has a slope equal to unity and is independent of the device dimensions.

A device with a highly conducting silicon nitride layer next to the silicon (I_1) and a less conducting layer of oxynitride on top (I_2) was fabricated with the initial dimension as shown in Fig. 9. The V_{FB} vs V_{appl} relation is plotted for several thicknesses of I_2 , obtained by step etching. The curves show a linear relation with the slope close to unity and independent of the device dimensions. The deviation from unity slope is probably due to decay of stored charge that takes place during the interval between the removal of the storing bias and the application of the bias that senses V_{FB} . The storage-bias relation was also obtained for another MI_2I_1S device in which the I_2 layer was pyrolytic silicon dioxide instead of oxynitride, as in the previous device. This sample also indicates that the approximation gives a good description of the device performance.

At this point the explanation for the saturation of V_{FB} with increasing V_{appl} is not clear. Also, the curves shown indicate the storage-bias relation for a negative applied voltage only. In the positive polarity, the V_{FB} vs V_{appl} relation is not linear, indicating that refinements are required for this simple approach.

VI FUTURE WORK

As has been demonstrated in this report, a novel application of a variable threshold transistor was possible. With further application studies, many more "untapped" uses may be uncovered. Since there is some linear dependence of V_T on the gate bias, many analog storage applications are feasible. Within the framework of this contract, however, no further work on adaptive circuits

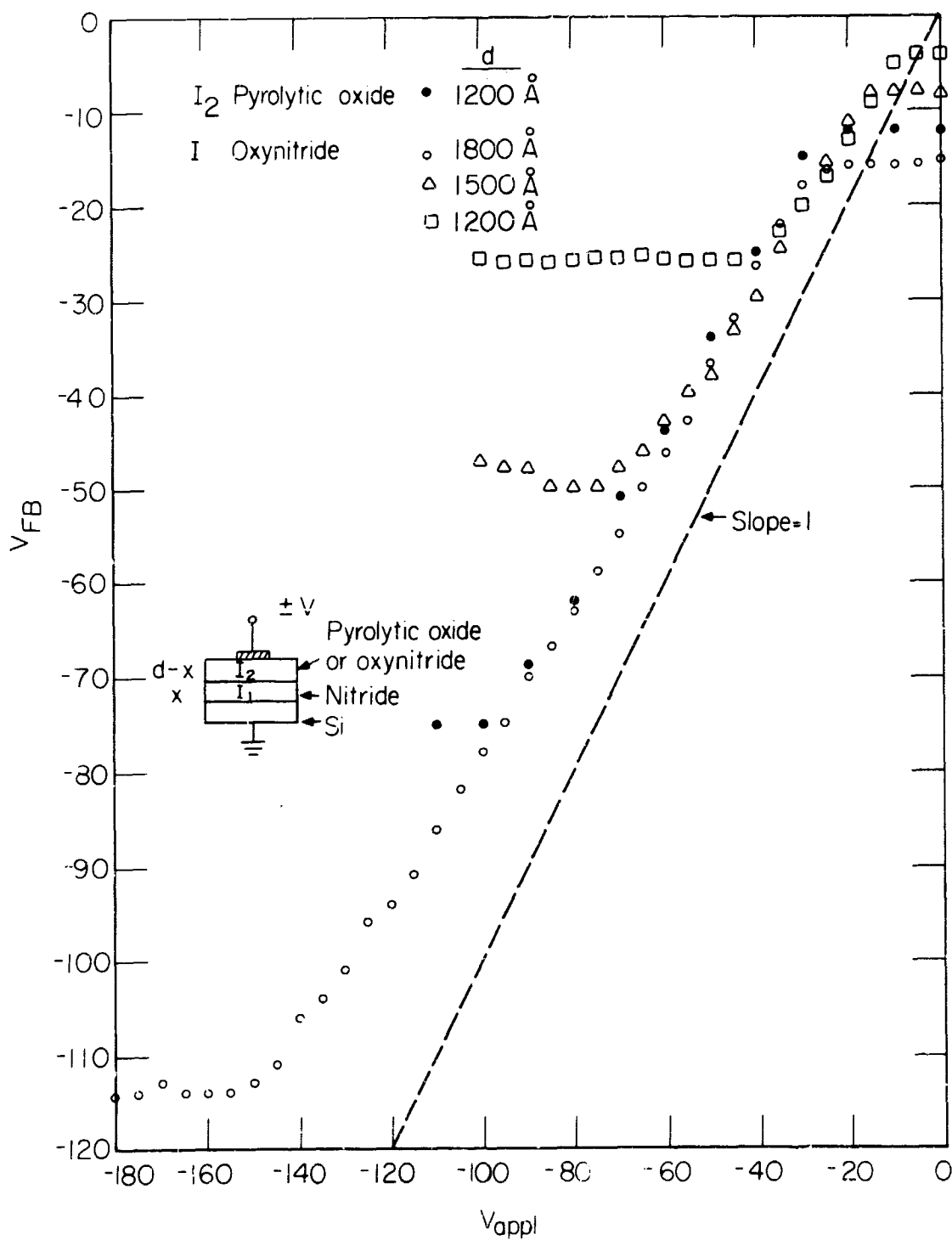


FIG. 9 V_{fb} vs V_{appl} .

is contemplated. The two layer model of the gate structure has been verified and it will be possible to incorporate an electrical equivalent circuit of the device in any general analysis. Work on a solution for the charge transient is being confirmed.

Since the device characterization depends on a knowledge of the device parameters, further work will be done on the control of processing parameters, such as layer thickness control and surface treatments prior to the deposition of both layers.

NEW TECHNOLOGY APPENDIX

After a diligent review of the work performed under this contract,
no new innovation, discovery, improvement or invention was made.